# REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188



1. REPORT DATE (DD-MM-YYYY): 16-12-2008

2. REPORT TYPE: Final Report

3. DATES COVERED (From - To): 17-Jul-2007 - 16-Apr-2008

4. TITLE AND SUBTITLE: Multi-modal, multi-signal data acquisition and processing based on compressive sensing

5a. CONTRACT NUMBER: W911NF-07-1-0502

5b. GRANT NUMBER:

5c. PROGRAM ELEMENT NUMBER: 611102

5d. PROJECT NUMBER:

5e. TASK NUMBER:

5f. WORK UNIT NUMBER:

6. AUTHORS: Richard G Baraniuk

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
William Marsh Rice University
William Marsh Rice University
6100 Main
Houston, TX 7725 -1182

8. PERFORMING ORGANIZATION REPORT NUMBER:

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S): ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S): 53332-CS-II.1

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The rapid growth of sensing and imaging technology combined with the need for near real-time action based on the sensed data has rendered automatic processing, understanding, and decision making vital to our national security. Our focus in this project is automatic target recognition based on imformation derived from families of related signals and images that are obtained from multiple imaging views or

15. SUBJECT TERMS
manifolds, ATR, compressive sensing, dimensionality reduction, classification, learning

| 16. SECURITY CLASSIFICATION OF: a. REPORT | b. ABSTRACT | c. THIS PAGE | 17. LIMITATION OF ABSTRACT | 15. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| U | U | U | SAR | | Richard Baraniuk |
| | | | | | 19b. TELEPHONE NUMBER |
| | | | | | 713-348-5132 |

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

## Report Title

Multi-modal, multi-signal data acquisition and processing based on compressive sensing

## ABSTRACT

The rapid growth of sensing and imaging technology combined with the need for near real-time action based on the sensed data has rendered automatic processing, understanding, and decision making vital to our national security. Our focus in this project is automatic target recognition based on imformation derived from families of related signals and images that are obtained from multiple imaging views or modalities. Key elements include random projections, dimensionality reduction techniques, and the new theory of compressive sensing. Result highlights include the development of new theory and algorithms for a compressive matched filter (the "smashed filter") and for a multimodal data fusion based on joint manifolds.

---

## List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

### (a) Papers published in peer-reviewed journals (N/A for none)

M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, “Single-Pixel Imaging via Compressive Sampling,” IEEE Signal Processing Magazine, special issue on Compressive Sampling, Vol. 25, Issue 2, pp. 83–91, March 2008.

R. G. Baraniuk, M. Davenport, R. A. DeVore, and M. B. Wakin, “A Simple Proof of the Restricted Isometry Property for Random Matrices” (aka “The Johnson-Lindenstrauss Lemma Meets Compressed Sensing,”), Constructive Approximation, January 2008.

R. G. Baraniuk and M. B. Wakin, “Random Projections of Smooth Manifolds,” Foundations of Computational Mathematics, December 2007.

**Number of Papers published in peer-reviewed journals:** 3.00

---

### (b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

**Number of Papers published in non peer-reviewed journals:** 0.00

---

### (c) Presentations

KEYNOTE AND PLENARY TALKS AT MAJOR CONFERENCE AND WORKSHOPS
2007/2008

“Compressive Sensing Theory and Applications,” IMA (UK) Conference on Mathematics of Signal Processing VIII, Cirencester, UK.

“Exploiting Sparsity through Compressive Sampling,” Workshop on Sparsity and its Application to Large Inverse Problems, Robinson College, Cambridge University.

“Distributed Compressive Sensing,” Sensor, Signal, and Information Processing Workshop, Sedona, AZ.

“Compressive Sensing,Wavelets, and Sparsity,” SPIE Defense + Security(acceptance speech for SPIE Wavelet Pioneer Award), Orlando.

“Compressive Signal Processing,” 42nd Conference on Information Sciences and Systems (CISS), Princeton.

“Compressive Signal Processing,” IAM-PIMS-MITACS Distinguished Colloquium Series, UBC, Vancouver.

“Compressive Sensing: A New Framework for Computational Data Acquisition,” CAMSAP (Computational Advances in Multi-Sensor Adaptive Processing, St. Thomas, Virgin Islands.

“Compressive Imaging for Vision Applications,” National Instruments Vision Summit, Austin.

“Compressive Detection and Estimation via Smashed Filtering,” AMS 2007 von Neumann Symposium on Sparse Representation and High-Dimensional Geometry, Snowbird.

“Compressive Imaging,” International Imaging Industry Association (I3A) 61st Annual Conference, Denver.

**Number of Presentations:** 10.00

## Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

**Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):** 0

## Peer-Reviewed Conference Proceeding publications (other than abstracts):

V. Cevher, C. Hegde, M. Duarte, and R. G. Baraniuk, “Sparse Signal Recovery Using Markov Random Fields,” Neural Information Processing Systems (NIPS), Vancouver, December 2008.

V. Cevher, P. Boufounos, M. Duarte, and R. G. Baraniuk, “Space Cutting for Distributed Localization,” 42nd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, October, 2008.

T. Ragheb, J. N. Laska, R. G. Baraniuk, and Y. Massoud, “Implementation of a Random Demodulation Based Compressive Analog-to-Digital Converter,” IEEE Midwest Symposium on Circuits and Systems, Knoxville, August 2008.

V. Cevher and R. G. Baraniuk, “Compressive Sensing for Sensor Calibration,” Fifth IEEE Workshop on Sensor Array and Multi-Channel Signal Processing (SAM), Darmstadt, Germany, July 2008.

V. Cevher, M. Duarte, and R. G. Baraniuk, “Distributed Target Localization via Spatial Sparsity,” 16th European Signal Processing Conference—EUSIPCO, Lausanne, August 2008.

P. Boufounos and R. G. Baraniuk, “Reconstructing Sparse Signals from their Zero Crossings,” IEEE International Conference on Acoustics, Speech and Signal Processing – ICASSP’08, Las Vegas, May 2008.

M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, “Wavelet-domain Compressive Signal Reconstruction using a Hidden Markov Tree Model,” IEEE International Conference on Acoustics, Speech and Signal Processing – ICASSP’08, Las Vegas, May 2008.

S. Pfetsch, T. Ragheb, J. N. Laska, H. Nejati, A. Gilbert, M. Strauss, R. G. Baraniuk, and Y. Massoud, “A Hardware Prototype for Random-Sampling Based Sub-Nyquist Analog-to-Information Conversion,” IEEE International Symposium on Circuits and Systems (ISCAS), Seattle, May 2008.

W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, “A Single-Pixel Terahertz Camera,” Conference on Lasers and Electro-Optics (CLEO), San Jose, May 2008.

P. Boufounos and R. G. Baraniuk, “One-Bit Compressive Sensing,” Conference on Information Sciences and Systems(CISS), Princeton, March 2008.

C. Hegde, M. Davenport, M. Wakin, and R. G. Baraniuk, “Efficient Machine Learning using Random Projections,” Neural Information Processing Systems (NIPS) Workshop on Efficient Machine Learning, Whistler, December 2007.

M. Davenport, C. Hegde, M.Wakin, and R. G. Baraniuk, “Manifold-based Approaches for Improved Classification,” Neural Information Processing Systems (NIPS) Workshop on Topology Learning, Whistler, December 2007.

C. Hegde, M.Wakin, and R. G. Baraniuk, “Random Projections for Manifold Learning,” Neural Information Processing Systems (NIPS), Vancouver, December 2007.

C. J. Rozell, D. H. Johnson, R. G. Baraniuk, B. A. Olshausen, “Locally Competitive Algorithms for Sparse Approximation,” IEEE International Conference on Image Processing – ICIP-2007, San Antonio, TX, September 2007.

M. Duarte, M. Davenport, M. Wakin, J. Laska, D. Takhar, K. F. Kelly, R. G. Baraniuk, “Multiscale Random Projections for Compressive Classification,” IEEE International Conference on Image Processing – ICIP-2007, San Antonio, TX, September 2007.

M. Sheikh, R. G. Baraniuk, “Blind Error-free Detection of Transform-Domain Watermarks,” IEEE International Conference on Image Processing – ICIP-2007, San Antonio, TX, September 2007.

P. T. Boufounos, M. F. Duarte, and R. G. Baraniuk, “Sparse Signal Reconstruction from Noisy Compressive Measurements Using Cross Validation,” IEEE Statistical Signal Processing Workshop, Madison, WI, August, 2007.

M. Davenport, R. G. Baraniuk, and C. Scott, “Minimax support vector machines,” IEEE Statistical Signal Processing Workshop, Madison, WI, August, 2007.

M. Moravec, J. K. Romberg, R. G. Baraniuk, “Compressive Phase Retrieval,” Wavelets XII in SPIE International Symposium on Optical Science and Technology, San Diego, August 2007.

**Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):** 19

## (d) Manuscripts

R. G. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based Compressive Sensing," submitted to IEEE Transactions on Information Theory, 2008.

M. Davenport, R. G. Baraniuk, and C. Scott, "Tuning Support Vector Machines for Minimax and Neyman-Pearson Classification, submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008.

D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed Compressed Sensing," submitted to IEEE Transactions on Information Theory, 2005.

**Number of Manuscripts:** 3.00

**Number of Inventions:**

### Graduate Students

| NAME | PERCENT_SUPPORTED |
|---|---|
| Mark Davenport | 0.14 |
| Marco Duarte | 0.14 |
| Chinmay Hegde | 0.14 |
| Shriram Sarvotham | 0.14 |
| Mona Sheikh | 0.14 |
| Matthew Moravec | 0.14 |
| Jason Laska | 0.14 |
| **FTE Equivalent:** | **0.98** |
| **Total Number:** | **7** |

### Names of Post Doctorates

| NAME | PERCENT_SUPPORTED |
|---|---|
| Petros Boufounos | 0.13 |
| Volkan Cevher | 0.13 |
| **FTE Equivalent:** | **0.26** |
| **Total Number:** | **2** |

### Names of Faculty Supported

| NAME | PERCENT_SUPPORTED | National Academy Member |
|---|---|---|
| Richard Baraniuk | 0.13 | No |
| **FTE Equivalent:** | **0.13** | |
| **Total Number:** | **1** | |

### Names of Under Graduate students supported

| NAME | PERCENT_SUPPORTED |
|---|---|
| **FTE Equivalent:** | |
| **Total Number:** | |

## Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: ...... 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense ...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

## Names of Personnel receiving masters degrees

NAME
Mark Davenport

**Total Number:** 1

## Names of personnel receiving PHDs

NAME
Shriram Sarvotham
Raymond Wagner

**Total Number:** 2

## Names of other research staff

| NAME | PERCENT_SUPPORTED |
| --- | --- |
| **FTE Equivalent:** | |
| **Total Number:** | |

## Sub Contractors (DD882)

## Inventions (DD882)

# Final Progress Report
# ARO Grant W911NF–07–1–0502
# Multi-signal, multi-modal data acquisition and processing based on compressive sensing

*Richard Baraniuk*

Rice University

**Abstract**

In this final report, we report on our progress on ARO grant W911NF–07–1–0502.

The goal of the project was to develop compressive sensing and dimensionality reduction for manifold data. We investigated and developed efficient sampling and measurement schemes for manifold-modeled data to enable efficient new A/AiTR and fusion algorithms. Key elements included random projections, dimensionality reduction techniques, and the new theory of compressive sensing. We developed new methods for signal and image registration, reconstruction, fusion, classification, and detection from incomplete information based on new compressive matched filters, MACH filters, and pattern recognition techniques.

The research highlights detailed below are (i) a new smashed filter for dimensionally reduced classification and A/AiTR; (ii) new algorithms for machine and manifold learning based on random projections; (iii) joint manifold models and processing algorithms for multi-sensor and multi-modal data fusion.

## 1 Signal and Image Manifolds

The geometric notion of a low-dimensional manifold is a powerful tool for modeling high-dimensional data. Manifold models arise in cases where (*i*) a $K$-dimensional parameter $\theta$ can be identified that carries the relevant information about a signal and (*ii*) the signal $x_\theta \in \mathbb{R}^N$ changes as a continuous (typically nonlinear) function of these parameters. Some typical examples include a 1-D signal shifted by an unknown time delay (parameterized by the translation variable), a recording of a speech signal (parameterized by the underlying phonemes spoken by the speaker), and an image of a 3-D object at an unknown location captured from an unknown viewing angle (parameterized by the 3-D coordinates of the object and its roll, pitch, and yaw). In these and many other cases, the geometry of the signal class forms a nonlinear $K$-dimensional manifold in $\mathbb{R}^N$,

$$\mathcal{M} = \{x(\theta) : \theta \in \Theta\},$$

where $\Theta$ is the $K$-dimensional parameter space [1–3]. Low-dimensional manifolds have also been proposed as approximate models for nonparametric signal classes such as images of human faces or handwritten digits [4–6].

## 2 The Multiscale Smashed Filter for Compressive Classification

The theory of *compressive sensing* (CS) enables a sparse or compressible signal to be reconstructed from a small set of non-adaptive linear projections In some applications, this allows us to directly acquire a compressed representation of a signal, effectively combining the steps of sampling and compression. However, in many applications we are not interested in obtaining a precise reconstruction of the scene under view, but rather are only interested in making some kind of detection or classification decision. For instance, in target classification, we simply wish to identify the class to which our image belongs out of several possibilities.

We have developed an algorithm to support a new theory of *compressive classification* that enjoys the same benefits as compressive sensing. Our approach, based on a generalized maximum likelihood classifier (GMLC), is applicable to a wide variety of signal classification problems. Focusing on the problem of image classification, we use the fact that the set of images of a fixed scene with different imaging parameters (translation, scale, view angle, illumination, etc.) forms a low-dimensional, nonlinear manifold in the high-dimensional ambient image space. Exploiting recent results on random projections of manifolds, we design a pseudo-random measurement scheme and a new classification algorithm—the *smashed filter*—that can be viewed as a generalization of the classical matched filter to more challenging manifold settings. The smashed filter achieves high classification rates using only a small fraction of measurements compared to the dimensionality of the original images.

We begin by examining the problem of signal classification where the formation of the signal $x$ under each hypothesis depends on specific parameters; this results in a combined estimation and classification problem. Suppose a signal $x \in \mathbb{R}^N$ belongs to one of $P$ classes $\mathcal{C}_i, i = 1, \ldots, P$. We let hypothesis $\mathcal{H}_i$ signify that the signal $x$ belongs to class $\mathcal{C}_i$. For each class $\mathcal{C}_i$, the generation of the signal $x$ under hypothesis $\mathcal{H}_i$ is governed by a $K$-dimensional *manifold* $\mathcal{M}_i = \{f_i(\theta_i) : \theta_i \in \Theta_i\}$ embedded in the ambient signal space $\mathbb{R}^N$. Example parameters for image classification scenarios include the pose of the object in the scene, translation, rotation, scale, etc. We obtain noisy measurements of $x$, as in $y = x + \omega \in \mathbb{R}^N$, giving us a distribution $p(y|\theta_i, \mathcal{H}_i)$ for the measured signal $y$ under hypothesis $\mathcal{H}_i$ and parameter $\theta_i$.

In the case of two classes, where the optimal classifier is the likelihood ratio test, we can accommodate these unknown parameters through the use of the *generalized* likelihood ratio test. We will refer to the multi-class extension of this technique as the *generalized maximum likelihood classifier* (GMLC). The GMLC reduces to a maximum likelihood estimator (MLE) for the parameter $\widehat{\theta_i}$ for each hypothesis $\mathcal{H}_i$, followed by an standard MLC using the parameter estimates. We can interpret the GMLC geometrically: the MLE of the parameter $\theta_i$ under the AWGN model corresponds to finding the closest point on the manifold $\mathcal{M}_i$ to the observed signal $y$. Subsequently, the MLC can be interpreted as a "nearest-manifold" search from $y$ to each of the $\mathcal{M}_i$.

We now consider the same classification problem where each class corresponds to the presence of a known signal in noise, but instead of observing $x+\omega$ we observe $y = \Phi(x+\omega)$ where $\Phi \in \mathbb{R}^{M \times N}$, $M \leq N$. In this case, the GMLC is essentially unchanged: our classifier again reduces to a "nearest-manifold" classifier; the only significant difference is that the $P$ classes now correspond to the manifolds $\Phi\mathcal{M}_i \subset \mathbb{R}^M$, $i = 1, \ldots, P$.[1] We dub this GMLC the *smashed filter* [8, 9]. The performance of the smashed filter does not depend on any structure of the signals $s_i$. Rather, its performance depends on the stability of the dimensionality reduction of the manifold: if the

[1]Linear projection by $\Phi$ of a manifold $\mathcal{M} \in \mathbb{R}^N$ yields another manifold in $\Phi\mathcal{M} \in \mathbb{R}^M$ [7].

(a) Tank (b) School Bus (c) Truck

Figure 1: *Models used for smashed filter classification experiments.*

distances between the projected points of the manifold and the projected signal are not preserved, then the estimator performance will suffer. This issue becomes critical during the nearest-neighbor classification step.

Let us now more closely examine the random projection of one or more manifolds from a high-dimensional ambient space $\mathbb{R}^N$ to a lower-dimensional subspace $\mathbb{R}^M$. We have shown that this process actually preserves the essential structure of a *smooth* manifold, provided that a sufficient number $M$ of random projections are taken:

**Theorem 2.1** *[7] Let $\{\mathcal{M}_i\}_{i=1}^P$ be compact $K$-dimensional submanifolds of $\mathbb{R}^N$ having condition numbers $1/\tau_i$ and volumes $V_i$, respectively. Fix $0 < \epsilon < 1$ and $0 < \rho < 1$ and let*

$$V = \sum_i V_i \quad \text{and} \quad \tau = \min\left(\min_i \tau_i, \ \min_{i \neq j} \ \mathrm{dist}(\mathcal{M}_i, \mathcal{M}_j)\right).$$

*Let $\Phi$ be a random orthoprojector from $\mathbb{R}^N$ to $\mathbb{R}^M$ with*

$$M = O\left(\frac{K \log(NV\tau^{-1}\epsilon^{-1}) \log(1/\rho)}{\epsilon^2}\right). \tag{1}$$

*If $M \leq N$, then with probability at least $1 - \rho$ the following statement holds: For every pair of points $x_1, x_2 \in \cup_i \mathcal{M}_i$,*

$$(1 - \epsilon)\sqrt{\frac{M}{N}} \leq \frac{\|\Phi x_1 - \Phi x_2\|_2}{\|x_1 - x_2\|_2} \leq (1 + \epsilon)\sqrt{\frac{M}{N}}.$$

Theorem 2.1 ensures not only that distances between pairs of points on each manifold are well-preserved, but also that the distances between the $P$ manifolds themselves are all well-preserved. In a classification setting with a large number of possible classes, the *sublinear* growth in the required number of measurements is particularly attractive.

We now present experimental results that evaluate the smashed filter in a image target classification setting. We consider three classes, each for a different vehicle model: a tank, a school bus, and a truck (see Figure 1). All images are of size $128 \times 128$ pixels, hence $N = 16384$, and all measurement matrices are binary orthoprojectors obtained from a random number generator.

Our experiment is synthetic and concerns unknown shifts of a known image. In this case, $K = 2$ and we know the explicit structure of the three manifolds: each can be constructed by translating a reference image in the 2-D image plane. The shifted versions of each image, as well as their corresponding compressive measurements, were obtained synthetically using software. The first

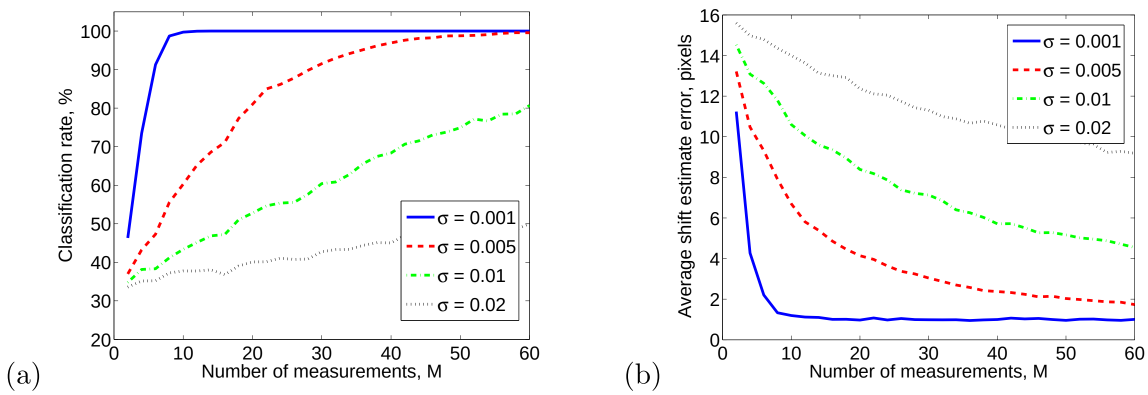


Figure 2: *Results for smashed filter image classification experiments. (a) Classification rates and (b) average estimation error for varying number of measurements $M$ and noise levels $\sigma$ for the image shift experiments. As $M$ increases, the distances between the manifolds increase as well, thus increasing the noise tolerance and enabling more accurate estimation and classification. Thus, the classification and estimation performances improve as $\sigma$ decreases and $M$ increases in all cases.*

step in implementing the smashed filter is thus to find the ML estimate of the shift for each class. This can be accomplished by simply calculating the distance between the observed $y$ and projections of all possible shifts of the image. For simplicity, we assume that the set of possible shifts is limited by a maximum shift of 16 pixels in any direction. After the shift estimate is obtained for each class, the GMLC selects the class whose estimate is closest to the observed image.

We performed classification experiments for different numbers of compressive measurements, varying from $M = 2$ to 60, and for different levels of additive Gaussian noise, with standard deviations $\sigma = 0.001, 0.005, 0.01$ and 0.02. For each setting, we executed 10000 iterations of the experiment, where we selected a testing point at random with a different noise realization at each iteration. The classification rates in Figure 2(b) show a clear dependence between the number of measurements and the classification rate. Furthermore, performance for a given $M$ degrades as the noise level increases; as expected the classifier becomes unreliable when the noise level becomes comparable to the minimum distance between the projected manifolds. In Figure 2(c), we see a similar relationship between the noise level and the error in the parameter estimate. These results verify that increasing the number of measurements improves the quality of the estimates; and the performance of the classifier is clearly dependent on the performance of the parameter estimator for the appropriate class.

# 3 Efficient Machine and Manifold Learning Using Random Projections

The unimpeded growth in the size of datasets generated by signal acquisition systems (e.g., sensor networks, 3D imaging systems) poses a significant challenge to machine learning algorithms. This effect – frequently referred to as the "curse of dimensionality" – usually forces an algorithm designer to sacrifice accuracy in order to make the problem computationally feasible.

Nevertheless, in many cases we can avoid this difficult decision. Suppose our dataset $X$ consists of points $x \in \mathbb{R}^N$. Often, points in $X$, although $N$-dimensional, can be described using a manifold model with only $K$ degrees of freedom, where $K \ll N$. A simple method to alleviate the curse of

dimensionality is as follows: compute a non-adaptive linear projection of the $N$-dimensional dataset into a *random* $M$-dimensional subspace of $\mathbb{R}^N$. In this case, the mapping $f$ can be represented as an $M \times N$ matrix $\Phi$ where the entries of $\Phi$ are independently drawn from a specified probability distribution. The simplicity of this dimensionality reduction procedure is striking; it is clear that the mapping is data independent, and the process of obtaining the image of any given data vector $x$ under the mapping $f$ is a stand-alone computation. In addition, the main result in [7] guarantees that, provided $M = O(K \log N)$, then there exists an $\epsilon \in (0, 1)$ such that

$$(1-\epsilon)\|x-y\| \leq \|\Phi x - \Phi y\| \leq (1+\epsilon)\|x-y\| \tag{2}$$

for all $x, y \in X$, where $\|.\|$ represents either the ambient 2-norm in $\mathbb{R}^N$, or the Riemannian distance on the manifold describing $X$. In other words, the distance between any pair of points is approximately preserved by the mapping $\Phi$.

This paves the way for an alternate method to solve high-dimensional machine learning problems. Let $\mathcal{L}$ denote some machine-learning algorithm tailored to the problem we wish to solve. Our claim is as follows: for a wide variety of machine-learning algorithms $\mathcal{L}$, the performance of $\mathcal{L}$ when given access to only a randomly projected (i.e. $M$-dimensional) version of $X$ is *essentially the same as* its performance on the original dataset $X$. The implications of this are significant; this implies that the machine is oblivious to whether it works with the original data, or with only a low-dimensional, easily obtainable representation. In other words, random projections can be used as a *universal, inexpensive* preprocessing step to almost any machine learning task. Thus, the random projections approach could lead to tremendous savings in processing and memory costs incurred during the learning process. In [10,11], the above claim is rigorously proved for two special cases: 1) when $\mathcal{L}$ is the Grassberger-Procaccia algorithm for estimating intrinsic dimension of a point cloud; 2) when $\mathcal{L}$ is the Isomap algorithm for nonlinear dimensionality reduction of Euclidean manifolds.

For illustration purposes, we test our theory on a common dataset (Figure 3) found in the literature on dimension estimation and manifold learning - the Stanford face database. The face database is a collection of 698 artificial snapshots of a face ($N = 64 \times 64 = 4096$) varying under 3 degrees of freedom: 2 angles for pose and 1 for lighting dimension. The images are therefore believed to reside on a 3D-manifold in an ambient space of dimension 4096. Random projections of each sample in the databases were obtained by computing the inner product of the image samples with an increasing number of rows of the random projection operator $\Phi$. We note that in the case of the face database, for $M > 40 \ll N$, the Isomap residual variance on the randomly projected points closely approximates the variance obtained with full image data. Also, the estimated GP dimension stabilizes to a value of 4, which is identical to the estimated GP dimension using the full image samples.

## 4 Joint Manifold Models for Distributed Data Processing

In some data acquisition scenarios, many observations of the same event are acquired simultaneously (e.g., from multiple cameras), resulting in the acquisition of multiple manifolds that share the same parameter space. For example, in a sensor network there may be many sensors observing a single event from different vantage points, but the underlying phenomenon can often be described by a single global parameter (such as the location of the objects of interest in a camera or microphone network). Similarly, when sensing a single phenomenon with multiple modalities, such as video

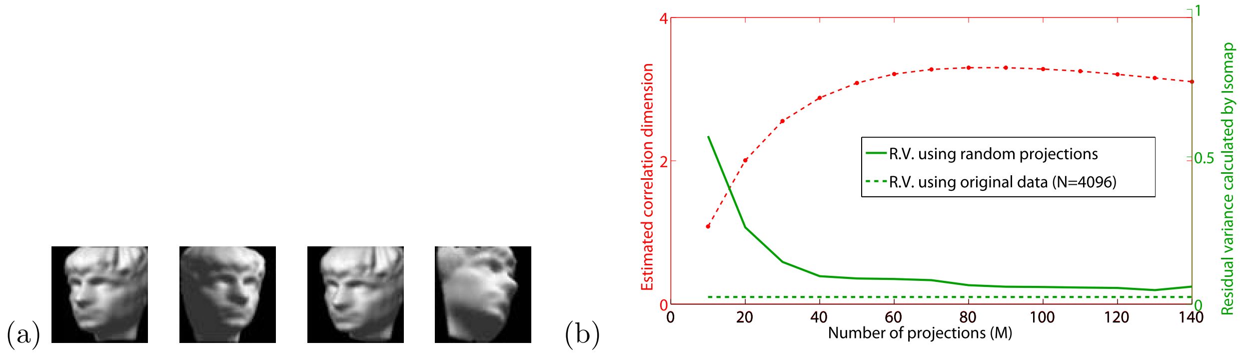


Figure 3: *(a) Images from the Stanford face database: ambient dimension $N = 4096$, intrinsic dimension $K = 3$ (b) Red curve indicates that just $M \approx 40 \ll N$ random projections carry sufficient information for accurate dimensionality estimation. The green curve indicates that again just $M \approx 40 \ll N$ random projections carry sufficient information to enable the nonlinear structure of the manifold to be learned accurately.*

and audio, the underlying phenomenon may again be described by a single modality-independent parametrization (such as what a speaker is saying in a recording of a person speaking). In such cases, we will see that it is advantageous to model the structure contained in the ensemble of manifolds rather than simply treating each manifold independently. Hence, we study a simple structure, the *joint manifold* on which we base our development of manifold-based learning and estimation algorithms that exploit this structure to improve their performance.

Specifically, let $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_J$ be an ensemble of $J$ manifolds of equal dimension $K$, and suppose that each manifold can be expressed as a function on a single parameter space: $\mathcal{M}_i = \{x_i(\theta) : \theta \in \Theta\}$. Denote the concatenation of all points for a single parameter value as $x(\theta) = [x_1(\theta)^T \dots x_J(\theta)^T]^T$. We define the *joint manifold* as the $K$-dimensional manifold resulting from this concatenation: $\mathcal{M}^* = \{x(\theta) : \theta \in \Theta\}$.

The advantage of considering the joint manifold, as opposed to the individual manifolds, lies in the fact that the joint manifold tends to be better behaved. Specifically, one can quantify this through the *condition number*, which controls local smoothness properties (such as curvature) as well as global properties (such as self-avoidance) of the manifold [12]. Intuitively speaking, the smaller the condition number, the more smooth the manifold. We show that the condition number of the joint manifold is less than the maximum of the condition numbers of the manifolds $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_J$. In practice we find that the joint manifold tends to be much better than this worst case analysis. In many cases we actually expect that the joint manifold will have a smaller condition number than the best of the component manifolds.

This property allows for the improvement of any manifold-based algorithm, including parameter estimation, detection, classification, and manifold learning. By running these algorithms on the joint manifold, rather than running them separately for each manifold, we can vastly improve their performance. However, when the ambient dimension of the manifolds $N$ or the number of manifolds $J$ in an ensemble is large, the dimensionality of the joint manifold– $JN$ – may become impossibly large to perform any meaningful computations. As a solution, we turn to the random projection approach: suppose we *linearly* project our joint $JN$-dimensional data into a random $M$-dimensional subspace. The results of [7] show that this process will preserve the essential properties of the manifold.

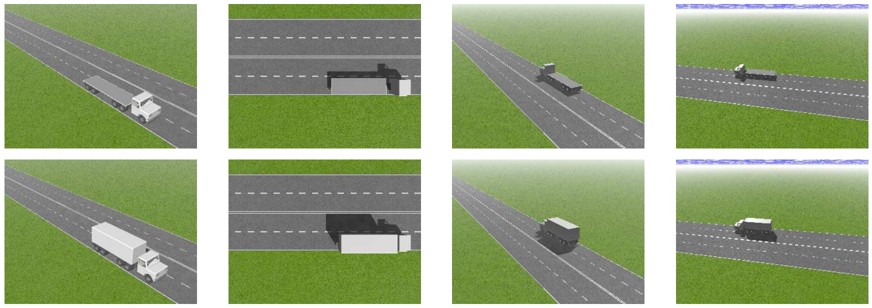

Figure 4: *Sample images of 2 different trucks from 2 different camera views. The motion of the truck along the road forms a 1-D manifold. The number of pixels in each camera image $N = 240 \times 360 = 76800$.*

In a distributed setting, we must also consider the cost involved in collecting and transmitting all of this data to a central node. Random projections also provide an elegant solution to this problem. Suppose the sensors are allowed to perform only local random projections of the observed signals prior to transmission. Fortunately, since the required measurements are linear, we can calculate *global* measurements of the JAM in a distributed fashion. Let each sensor obtain its measurements $z_i = \Phi_i x_i$, with the matrices $\Phi_i \in \mathbb{R}^{M \times N}, i = 1, \ldots, J$. Then, by defining the $M \times JN$ matrix $\Phi = [\Phi_1 \ldots \Phi_J]$, our global projections $z = \Phi x$ can be obtained by

$$\begin{aligned} z &= \Phi x \\ &= \Phi[x_1^T \ldots x_J^T]^T \\ &= [\Phi_1 \ldots \Phi_J][x_1^T \ldots x_J^T]^T \\ &= \Phi_1 x_1 + \ldots + \Phi_J x_J. \end{aligned}$$

Thus, the final measurement vector can be obtained by simply *adding independent random projections* of the signals acquired by the individual sensors.

As a stylized example, in this section we apply the random projections-based fusion algorithm to the task of moving vehicle classification. Our images are generated using POVRAY (*http://www.povray.org*), an open-source ray tracing software package.

We let a number $J$ of cameras, each with resolution $N$, observe the motion of a truck along a straight road. This forms an ensemble of $J$ 1-dimensional manifold in the image space $\mathbb{R}^N$, one formed from each camera; the joint manifold describing the truck's motion would also be one-dimensional in $\mathbb{R}^{JN}$. Our classification task consists of distinguishing between two types of trucks with minimal information processing. Example images from two different camera views for the two classes are shown in Table 4. In our experiment, we obtain 200 random projections from the 1D translational parameter space for three camera views. The resolution of each image is $N = 240 \times 320 = 76800$ pixels. The sample camera views suggest that some views make it easier to distinguish between the classes than others. For instance, the head-on view of the two trucks is very similar for most shift parameters, while the side view is appropriate for detecting the difference between the two classes of trucks.

The error curves of joint manifold classification as a function of the signal-to-noise ratio (SNR) are shown in Figure 5(a). It is clear that joint manifold classification does better than majority voting and is comparable in performance to the best camera. However, in the absence of prior

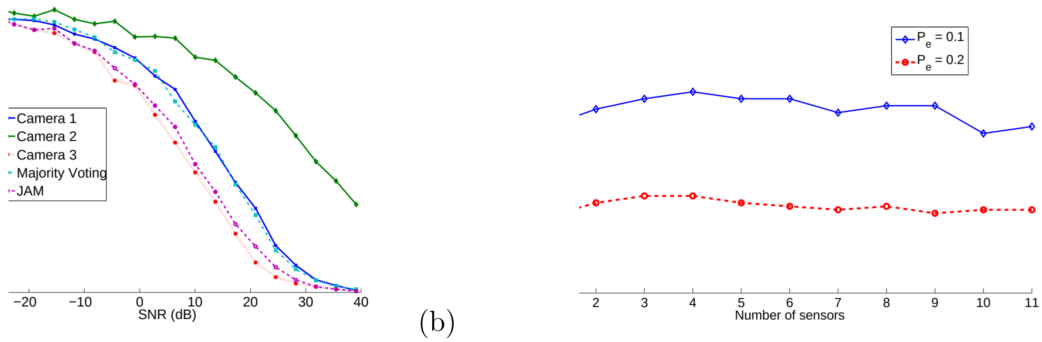


Figure 5: *(a) SNR vs. probability of error curves for individual cameras, JAM, majority voting. JAM performs better than majority voting and nearly as well as the best camera. (b) Number of sensors vs communication bandwidth. JAM-based fusion ensures that the total amount of transmitted information to the central unit remains approximately the same with increasing $J$.*

information regarding how well each camera truly performs, the best strategy for the central processor would be to fuse the data from *all* cameras. Thus, joint manifolds proves to be more effective than high-level fusion algorithms like majority voting.

A crucial feature of joint manifold data fusion is that with increasing number of sensors, the *total* communication bandwidth (proportional to the number of measurements per sensor $M$) does not need to increase rapidly. Figure 5(b) describes the variation of the number of fused measurements transmitted to the central decision unit by each sensor so as to obtain a given probability of error $P_e$ under an SNR value of 10 dB per measurement. The plots are approximately flat; thus, this demonstrates the utility of joint manifold fusion techniques in scenarios where bandwidth is a severe constraint.

## References

[1] D. L. Donoho and C. Grimes. Image manifolds which are isometric to Euclidean space. *J. Math. Imaging and Computer Vision*, 23(1), July 2005.

[2] C. Grimes. *New methods in nonlinear dimensionality reduction*. PhD thesis, Department of Statistics, Stanford University, 2003.

[3] M. B. Wakin, D. L. Donoho, H. Choi, and R. G. Baraniuk. The multiscale structure of non-differentiable image manifolds. In *Proc. Wavelets XI at SPIE Optics and Photonics*, San Diego, August 2005. SPIE.

[4] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1), 1991.

[5] G. E. Hinton, P. Dayan, and M. Revow. Modelling the manifolds of images of handwritten digits. *IEEE Trans. Neural Networks*, 8(1), 1997.

[6] D. S. Broomhead and M. J. Kirby. The Whitney Reduction Network: A method for computing autoassociative graphs. *Neural Computation*, 13:2595–2616, 2001.

[7] R. G. Baraniuk and M. B. Wakin. Random projections of smooth manifolds. 2007. To appear in *Foundations of Computational Mathematics.*

[8] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk. The smashed filter for compressive classification and target recognition. In *Proc. IS&T/SPIE Symposium on Electronic Imaging: Computational Imaging*, volume 6498, San Jose, CA, Jan. 2007.

[9] M. F. Duarte, M. A. Davenport, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk. Multiscale random projections for compressive classification. In *IEEE International Conference on Image Processing (ICIP)*, pages VI–161–164, San Antonio, TX, Sept. 2007.

[10] C. Hegde, M.B. Wakin, and R.G. Baraniuk. Random projections for manifold learning. In *Neural Information Processing Systems (NIPS)*, 2007.

[11] C. Hegde, M. B. Wakin, and R. G. Baraniuk. Random projections for manifold learning - proofs and analysis. Technical Report TREE, Rice University ECE Department, Houston, TX, 2007.

[12] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with confidence from random samples. Technical Report TR-2004-08, University of Chicago, 2004.